



# Ensuring data access, transparency, and preservation: mandatory data deposition for *Behavioral Ecology and Sociobiology*

Theo C. M. Bakker<sup>1</sup> · James F. A. Traniello<sup>2</sup>

© The Author(s) 2020

Universal data sharing allows access to data sets supporting published research and enables the use of data for additional applications. It is important for quality control to ensure compliance with best practices for publication and scientific ethics and reliability of data storage. Data deposition in repositories is obligatory for publications involving molecular data (e.g., nucleotide sequence data, protein structural data) and for all data in publications in the life sciences and earth sciences in the high-impact journals *Nature*, *Science*, *Proceedings of the National Academy of Sciences USA*, and *Proceedings of the Royal Society B* (see <https://www.nature.com/sdata/policies/repositories> for recommended repositories for *Nature*). Data sharing allows the preservation and reuse of data when ethical (Duke and Porter 2013; Mills et al. 2015). Data not stored in a repository may be lost in the future: in a study of more than 500 papers containing morphological data, each annual increase in article age was found to decrease the odds of the data set remaining extant by 17% (Vines et al. 2013). Data sharing also facilitates the detection of irregularities or anomalies in raw data that may lead to corrections or retractions of published papers in behavioral ecology (e.g., Keiser et al. 2020; Laskowski et al. 2020a, b; Proceedings B Editorial Team 2020). Mandatory data sharing can thus strengthen journal credibility and stature and generally advance scientific interests (e.g., Piwowar and Vision 2013). Data upload, however, is uncommon in the publication of behavioral research (Caetano and Aisenberg

2014). Mandatory requirements for making data available are increasing (Setchell et al. 2016; Sim et al. 2020).

*Behavioral Ecology and Sociobiology* will now require data deposition and strongly recommends to supplement source codes for publications in the journal to strengthen our policy of promoting the transparency and reproducibility of behavioral science by making data sets accessible (see Instructions for Authors for recommended repositories). This change in policy will ultimately benefit authors, as papers with publicly available datasets have higher citation rates than similar studies that do not post data (Piwowar and Vision 2013). Sharing long-term data may come with serious costs to the principal investigator; Mills et al. (2015) therefore suggest that “journals should have a rule that no paper is considered where the data users have not corresponded with the data owners and included appropriate acknowledgment of the source of the data within the paper.”

Data for articles published in *Behavioral Ecology and Sociobiology* should be archived properly; recommendations are provided by Roche et al. (2015) and are available on upload service websites. Error should be minimized, or the full benefits of data availability will not be realized. For example, in a study of 100 datasets in *Dryad* from non-molecular evolutionary and/or ecological publications in leading journals, 56% were incomplete, and 64% were archived in formats that partially or entirely prevented reuse (Roche et al. 2015). The most common problems with these data sets were inadequate metadata, the use of inadequate file formats, and failure to archive raw data. Proper data archiving has a high benefit, as completeness and reusability scores are strongly correlated (Roche et al. 2015).

Repositories may be institutional, national, or global. *Dryad*, *Drum*, and *EASY* have been identified as suitable repositories for general scientific data and have a number of favorable characteristics such as availability of guidelines for upload and storage and long-term preservation (Banzi et al. 2019). Sequence data must be deposited in disciplinary repositories. Institutional repositories may be able to better guarantee the ethical use of shared long-term data (Mills et al. 2015).

✉ Theo C. M. Bakker  
tbakker@evolution.uni-bonn.de

<sup>1</sup> Institute for Evolutionary Biology and Ecology, University of Bonn, An der Immenburg 1, 53121 Bonn, Germany

<sup>2</sup> Department of Biology, Boston University, 5 Cummington Mall, Boston, MA 02215, USA

**Acknowledgments** We thank Rebecca Grant for comments on an earlier draft.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Banzi R, Canham S, Kuchinke W, Krleza-Jeric K, Demotes-Mainard J, Ohmann C (2019) Evaluation of repositories for sharing individual-participant data from clinical studies. *Trials* 20:169
- Caetano DS, Aisenberg A (2014) Forgotten treasures: the fate of data in animal behaviour studies. *Anim Behav* 98:1–5
- Duke CS, Porter JH (2013) The ethics of data sharing and reuse in biology. *BioScience* 63:483–489
- Keiser CN, Sweeney K, Cusack B, Armagost F, Pruitt JN (2020) Corrigendum to: “predator and prey activity levels jointly influence the outcome of long-term foraging bouts” (2013). *Behav Ecol* 24: 1205–1210. <https://doi.org/10.1093/beheco/art052> (published online, doi:10.1093/beheco/araa038)
- Laskowski KL, Modlmeier AP, DeMarco AE, Coleman A, Zhao K, Brittingham HA, McDermott DR, Pruitt JN (2020a) Retraction: persistent social interactions beget more pronounced personalities in a desert-dwelling social spider. *Biol Lett* 16:20200062
- Laskowski KL, Montiglio P-O, Pruitt JN (2020b) Retraction. *Am Nat* 195:303
- Mills JA, Teplitsky C, Arroyo B, Charmantier A, Becker PH, Birkhead TR, Bize P, Blumstein DT, Bonenfant C, Boutin S, Bushuev A, Cam E, Cockburn A, Côté SD, Coulson JC, Daunt F, Dingemanse NJ, Doligez B, Drummond H, Espie RHM, Festa-Bianchet M, Frentiu F, Fitzpatrick JW, Furness RW, Garant D, Gauthier G, Grant PR, Griesser M, Gustafsson L, Hansson B, Harris MP, Jiguet F, Kjellander P, Korpimäki E, Krebs CJ, Lens L, Linnell JDC, Low M, McAdam A, Margalida A, Merilä J, Møller AP, Nakagawa S, Nilsson JÅ, Nisbet ICT, van Noordwijk AJ, Oro D, Pärt T, Pelletier F, Potti J, Pujol B, Réale D, Rockwell RF, Ropert-Coudert Y, Roulin A, Sedinger JS, Swenson JE, Thébaud C, Visser ME, Wanless S, Westneat DF, Wilson AJ, Zedrosser A (2015) Archiving primary data: solutions for long-term studies. *Trends Ecol Evol* 30:581–589
- Piwowar HA, Vision TJ (2013) Data reuse and the open data citation advantage. *PeerJ* 1:e175
- Proceedings B Editorial team (2020) Expression of concern: individual personalities shape task differentiation in a social spider. *Proc R Soc B* 287:20200846
- Roche DG, Kruuk LEB, Lanfear R, Binning SA (2015) Public data archiving in ecology and evolution: how well are we doing? *PLoS Biol* 13:e1002295
- Setchell JM, Fernandez-Duque E, Higham JP, Rothman JM, Schülke O (2016) Editorial: changes and clarifications to the policies of the *International Journal of Primatology* to promote transparency and open communication. *Int J Primatol* 37:617–627
- Sim I, Stebbins M, Bierer BE, Butte AJ, Drazen J, Dzau V, Hernandez AF, Krumholz HM, Lo B, Munos B, Perakslis E, Rockhold F, Ross JS, Terry SF, Yamamoto KR, Zarin DA, Li R (2020) Time for NIH to lead on data sharing a draft policy is generally supportive but should start mandating data sharing. *Science* 367:1308–1309
- Vines TH, Albert AYK, Andrew RL, Débarre F, Bock DG, Franklin MT, Gilbert KJ, Moore J-S, Renaut S, Rennison DJ (2013) The availability of research data declines rapidly with article age. *Curr Biol* 24: 94–97

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.